

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings of claims in the application. Applicant has submitted a new complete claim set showing any marked up claims with insertions indicated by underlining and deletions indicated by strikeouts and/or double bracketing.

1. (Previously Presented) A system for performing client-centric load balancing of multiple globally-dispersed servers, the servers being accessed by clients connecting through an ISP having a domain name server (DNS-ISP), the servers further having an authoritative domain name server (DNS-A) associated therewith, the system comprising:

a first domain name server deployed on an Internet backbone (DNS-B); and
a plurality of load balancing domain name servers (DNS-LBs) deployed in a physical proximity from which the actual network latency of the clients may be measured, the DNS-LBs having stored therein IP address information of the multiple globally-dispersed servers to be load balanced, the DNS-LBs each sending mapping information to the DNS-B relating the DNS-LB's IP address to an IP address of the DNS-ISP to which the DNS-LB is in a physical proximity from which the actual network latency of the clients may be measured, the DNS-LBs determining performance characteristics of each of the multiple globally-dispersed servers.

2. (Original) The system of claim 1, wherein the DNS-B stores the mapping information for the plurality of DNS-LBs to forward IP address queries to one of the DNS-LBs closest to the DNS-ISP from which the IP address query originated, and wherein the DNS-LB closest to the DNS-ISP returns the IP address to the DNS-ISP of the server having the best performance characteristics.

3. (Original) The system of claim 1, wherein the DNS-B stores the mapping information for the plurality of DNS-LBs to forward IP address queries to one of the DNS-LBs closest to the DNS-ISP from which the IP address query originated, and wherein the DNS-LB closest to the DNS-ISP returns the IP address of the DNS-LB to the DNS-ISP.
4. (Original) The system of claim 1, wherein the DNS-B provides its IP address information to the DNS-A to enable the DNS-A to forward IP address queries to the DNS-B.
5. (Original) The system of claim 4, wherein the DNS-B receives IP address information from the DNS-A for the servers to be load balanced.
6. (Original) The system of claim 1, wherein the DNS-LB is a client of the DNS-ISP.
7. (Original) The system of claim 1, further comprising a DNS-B deployed on each Internet backbone, and wherein each DNS-B contains the mapping information for all of the DNS-LBs stored therein.
8. (Original) The system of claim 1, wherein the DNS-LB transmits updated mapping information upon a change of an IP address of the DNS-ISP.
9. (Original) The system of claim 1, wherein each of the DNS-LBs transmit performance information of the servers to the DNS-B, and wherein the DNS-B utilizes the mapping information to determine the proper DNS-LB performance information to utilize to select the IP address of the server having the best performance characteristics to return to the DNS-ISP from which an IP address query originated.

10. (Previously Presented) A method of performing client-centric load balancing of multiple globally-dispersed servers, the servers being accessed by clients connecting through an ISP having a domain name server (DNS-ISP), the servers further having an authoritative domain name server (DNS-A) associated therewith, the method comprising the steps of:

receiving IP address information from the DNS-A for the servers to be load balanced;
providing the IP address information to a plurality of load balancing domain name servers (DNS-LB);

receiving mapping information associating DNS-ISP IP address information to IP address information of a DNS-LB located in a physical proximity from which the actual network latency of the clients may be measured, the DNS-LB capable of determining server performance from a location physically proximate to the ISP's point of presence; and

referring address inquiries from a DNS-ISP to a physically proximate DNS-LB in accordance with the mapping information.

11. (Original) A computer-readable medium having computer executable-instructions for performing the steps of claim 10.

12. (Previously Presented) A method of performing client-centric load balancing of multiple globally-dispersed servers, the servers being accessed by clients connecting through an ISP having a domain name server (DNS-ISP), the servers further having an authoritative domain name server (DNS-A) associated therewith, the method comprising the steps of:

obtaining, by a load balancing domain name server (DNS-LB), IP address information for a DNS-ISP, the DNS-LB located in a physical proximity from which the actual network latency of the clients may be measured;

providing a mapping of an IP address of the DNS-LB to the IP address information of the DNS-ISP to an external domain name server;
receiving IP address information for the servers;
monitoring performance of the servers at the received IP addresses; and
providing at least one IP address for a server in response to a name query selected based on the monitoring step.

13. (Original) The method of claim 12, further comprising the steps of:
detecting a change in the DNS-ISP IP address; and
updating the mapping of the IP address of the DNS-LB to the IP address information of the DNS-ISP to the external domain name server.

14. (Original) The method of claim 12, further comprising the steps of
receiving selection criteria for the selection of an IP address;
receiving a name query from the DNS-ISP; and
wherein the step of providing at least one IP address for a server in response to a name query selected based on the monitoring step further comprises the step of providing at least one IP address for a server in response to a name query selected based on the monitoring step and on the selection criteria.

15. (Original) A computer-readable medium having computer-executable instructions for performing the steps of claim 12.

16-19. (Canceled).

20. (Previously Presented) A method of performing client-centric load balancing of multiple globally-dispersed servers, the servers being accessed by clients connecting through Internet service providers (ISPs) at a point of presence (POP), each ISP having a domain name server (DNS-ISP), the servers further having an authoritative domain name server (DNS-A) associated therewith containing information regarding the IP addresses of the servers, the method comprising the steps of:

- deploying a first plurality of load balancing domain name servers (DNS-LBs) in a physical proximity from which the actual network latency of the clients connecting to the ISP POPs may be measured;

- deploying a second plurality of second level domain name servers (DNS-Bs) on the Internet backbones and regional provides;

- communicating IP address information for the DNS-Bs to the DNS-As to enable the DNS-As to refer name queries to the DNS-Bs;

- providing, by the DNS-LBs to the DNS-B, mapping information associating an IP address of the DNS-LB to the DNS-ISP to an IP address of the DNS-ISP to enable the DNS-B to refer name queries from a DNS-ISP to DNS-LB to the DNS-ISP; and

- communicating IP address information of the servers to the DNS-LBs;

- monitoring, by the DNS-LBs, performance of the servers; and

- providing, by the DNS-LB in response to a query from the DNS-ISP, the IP address of a server based on the step of monitoring.

21. (Cancelled)

22. (Cancelled)

23. (Currently Amended) A method for load balancing content servers, each of the content servers associated with a domain name, the method comprising:

receiving a request to resolve the domain name from an ISP DNS server;
identifying at least one load balancing server from a group of load balancing servers, the identified load balancing server situated at a physical proximity from which the actual network latency of a client connecting to the ISP DNS server may be measured; and
sending the IP address of the identified load balancing server to the ISP DNS server, the identified load balancing server configured to determine at least one of the content servers with certain characteristics relative to the location and to resolve the domain name with an IP address associated with the determined content server.

24. (Previously Presented) The method as recited in claim 23, wherein the certain characteristics include load level, availability, network latency, or network cost.

25. (Previously Presented) The method as recited in claim 23, wherein the identified load balancing server is situated closest to the ISP DNS server among the group of load balancing servers.

26. (Previously Presented) A system for load balancing a group of content servers located at multiple sites, the group of content servers associated with a domain name, each content server configured to interact with clients through ISP DNS servers associated with the clients, the system comprising:

load balancing servers configured to resolve the domain name with an IP address associated with at least one of the content servers, each load balancing server situated at a physical proximity from which the actual network latency of a client connecting to at least one of the ISP DNS servers may be measured, each load balancing server also configured to monitor

the content servers and to resolve the domain name with the IP address of at least one of the content servers with certain characteristics relative to the location of the load balancing server;

a referral server configured to receive requests to resolve the domain name from the ISP DNS servers, the referral server configured to respond to each request by determining at least one of the load balancing servers from which the request was received and referring the ISP DNS server to the determined load balancing server;

wherein requests to resolve the domain name from each ISP DNS server are responded by a load balancing server, and wherein each load balancing server responds to each request by determining at least one of the content servers that has certain characteristics relative to the location and by resolving the domain name in the request with the IP address of the determined content server.

27. (Previously Presented) The system as recited in claim 26, wherein the certain characteristics include load level, availability, network latency, or network cost.

28. (New) A method performed by a DNS-ISP that provides DNS service for a plurality of clients, the method comprising:

receiving from a client a lookup request requesting an IP address for a web server having a hostname specified by the request;

in response to receiving the client request, determining that an IP address for the requested hostname is needed and in response issuing a DNS query for the hostname;

in response to issuing the DNS query for the hostname by the DNS-ISP, receiving a redirection to an authoritative DNS server (DNS-A) that corresponds to the hostname, the referral providing an IP address of a domain name service load-balancing server (DNS-LB) and causing the DNS-ISP to query the DNS-LB for the IP address of the hostname in the client request;

in response to querying the DNS-LB, receiving the IP address from the DNS-LB, where the IP address was selected by the DNS-LB based on transmissions from the DNS-LB to the IP address that measure network latency from the DNS-LB to the IP address of the server; and sending the IP address received from the DNS-LB to the client that sent the lookup request for the web server having the hostname specified by the request.

29. (New) A method performed by a load-balancing domain name service server (DNS-LB) that provides load balancing service for a plurality of clients, the method comprising:

receiving at the DNS-LB a request received by and forwarded from a domain name service server of an Internet Service Provider (DNS-ISP), the request having been sent by a client of the DNS-ISP, the request containing a hostname corresponding to a plurality of IP addresses of servers serving web content being requested by the request, where the request was forwarded by the DNS-ISP when the DNS-ISP: determined that an IP address for the hostname was not cached at the DNS-ISP and obtained the IP address of the DNS-LB by issuing a DNS query for the hostname;

measuring network latency from the DNS-LB to the one or more IP addresses corresponding to the hostname in the request by periodically sending communications from the DNS-LB to the one or more IP addresses;

in response to receiving the request of the client at the DNS-LB, selecting an IP address from among the IP addresses corresponding to the hostname, where the IP address is selected based on the measuring of network latency; and

returning to the client the selected IP address to be used by the client to send its request for web content to the IP address.